09/920467

Part of Paper # 13
Figures not included

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
15 February 2001 (15.02.2001)

PCT

(10) International Publication Number
WO 01/11074 A2

(51) International Patent Classification[7]: C12Q

(21) International Application Number: PCT/US00/21223

(22) International Filing Date: 3 August 2000 (03.08.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/369,364 6 August 1999 (06.08.1999) US

(71) Applicant (*for all designated States except US*): THE CLEVELAND CLINIC FOUNDATION [US/US]; 9500 Euclid Avenue, Cleveland, OH 44195 (US).

(71) Applicants and
(72) Inventors: APTE, Suneel, S. [IN/US]; 19917 Malvern Road, Shaker Hts., OH 44122 (US). HURSKAINEN, Tina, L. [FI/FI]; Onnelantie 3A, FIN-90230 Oulu (FI). HIROHATA, Satoshi [JP/US]; 6809 Mayfield Road, #859, Mayfield Hts., OH 44124 (US).

(74) Agent: DOCHERTY, Pamela, A.; Calfee, Halter & Griswold LLP, Suite 1400, 800 Superior Avenue, Cleveland, OH 44114 (US).

(81) Designated States (*national*): AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

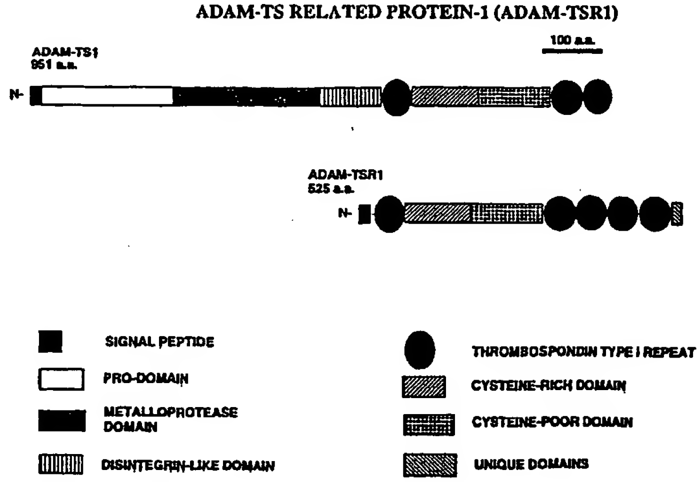
(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: NUCLEIC ACIDS ENCODING ZINC METALLOPROTEASES

ADAM-TS RELATED PROTEIN-1 (ADAM-TSR1)



(57) Abstract: Isolated mammalian proteins having disintegrin-like and metalloprotease domains with thrombospondin type I motifs, i.e., ADAMTS proteins, are provided. The proteins are ADAMTS-5, ADAMTS-6, ADAMTS-7, ADAMTS-8, ADAMTS-9 and ADAMTS-10, collectively referred to as "ADAMTS-N". The present invention also provides isolated polynucleotides which encode an ADAMTS-N protein or a variant thereof, polynucleotide sequences complementary to such polynucleotides, vectors containing such polynucleotides, and host cells transformed or transfected with such vectors. The present invention also relates to antibodies which are immunospecific for one or more of the ADAMTS-N proteins. The present invention also relates to a protein referred to hereinafter as ADAMTS-R1 (ADAM-TS Related protein-1) and the polynucleotides which encode such protein.

WO 01/11074 A2

NUCLEIC ACIDS ENCODING ZINC METALLOPROTEASES

Background of the Invention

This invention relates to isolated nucleic acid -molecules which encode proteins belonging to a zinc metalloprotease family. The zinc metalloproteases have been implicated in a variety of diseases and development disorders that involve* enhanced or depressed proteolysis of components of the extracellular matrix, receptors, or other extracellular molecules.

More particularly, the invention relates to isolated nucleic acid molecules encoding proteins belonging to a subfamily of zinc metalloproteases referred to as "ADAMTS", an abbreviation for A Disintegrin-like And Metalloprotease domain with ThromboSpondin type I motifs. Proteins in the ADAMTS subfamily all possess a Zn protease catalytic site consensus sequence (HEXXH+H), which suggests an intact catalytic activity for each of these proteins. The ADAMTS proteins also have putative N-terminal signal peptides and lack transmembrane domains, which suggests that the proteins in this subfamily are secreted. The proteins in the ADAMTS subfamily also possess at least one thrombospondin type (TSP1) motif, which suggests a binding of these proteins to components of the extracellular matrix (ECM) or to cell surface components.

Members of the ADAMTS subfamily of proteins are ADAMTS-1, ADAMTS-2, ADAMTS-3, and ADAMTS-4. ADAMTS-1 protein is selectively expressed in colon 26 adenocarcinoma cachexigenic sublines. ADAMTS-1 mRNA is induced by the inflammatory cytokine interleukin-1 in vitro and by intravenous administration of lipopolysaccharide in vivo. Thus, the ADAMTS-1 protein is believed to play a role in tumor cachexia and inflammation.

The ADAMTS-2 protein is also known as procollagen I/H amino-propetide processing enzyme r PCINP. The ADAMTS-2 protein catalyzes

cleavage of native triple-helical procollagen I and procollagen II. The ADAMTS-2 protein also has an affinity for collagen XIV. Lack of the ADAMTS-2 protein is known to cause dermatosparaxis in cattle, or Ehlers-Danlos syndrome type VIIC (EDS-VIIC) in humans. EDS-VIIC is characterized clinically by severe skin fragility, and biochemically by the presence in skin of procollagen which is incompletely processed at the amino terminus. Thus, it is believed that the ADAMTS-2 protein plays a role in processing of procollagen to mature collagen, an essential step for correct assembly of collagen into collagen fibrils. The ADAMTS-3 protein is similar in sequence to ADAMTS-2 and may have similar function.

The ADAMTS-4 protein catalyzes cleavage of the core protein of the extracellular matrix proteoglycan, aggrecan. Aggrecan degradation is an important factor in the erosion of articular cartilage in arthritic disease. Aggrecan fragments have been identified in cultures undergoing cartilage matrix degradation and in arthritic synovial fluids. Therefore, overexpression or activation 10 of ADAMTS-4 protein may be related to both inflammatory and non-inflammatory arthritis.

On the basis of the structure, location, and the demonstrated proteolytic activity of ADAMTS proteins 1-4, it is expected that other members of the ADAMTS subfamily play a role in the cleavage of proteoglycan core proteins that are found in the extracellular matrix, such as, for example, versican, brevican, neuracan, NG-2, aggrecan, as well as molecules such as collagen. It is also expected that other members of the ADAMTS subfamily play a role in embryogenesis, implantation of a fertilized egg, angiogenesis, arthritic degradation of cartilage, inflammation, nerve regeneration, tumor growth, and metastases.

Thus, it is desirable to have other members of the ADAMTS

subfamily of proteins, the nucleic acids that encode such proteins, and antibodies that are specific for such proteins. Such molecules are useful research tools for studying development of the extracellular matrix during embryogenesis and fetal development, and for studying disorders or diseases that are characterized by improper development of the extracellular matrix or enhanced or reduced destruction of the extracellular matrix. Such molecules, particularly the nucleic acids and the antibodies, are also useful tools for diagnosing such diseases or for monitoring the efficacy of therapeutic agents that have been developed to treat such diseases.

Summary of the Invention

The present invention provides novel, isolated, and substantially purified proteins having the characteristics of an ADAMTS protein. The novel proteins are referred to hereinafter individually as "ADAMTS-5", "ADAMTS-6", "ADAMTS-7", "ADAMTS-8", "ADAMTS-9" and "ADAMTS-10", and collectively as "ADAMTS-N". In one embodiment, the ADAMTS-5 protein is a mature mouse protein which comprises amino acid 231 through amino acid 930 of the sequence set forth set forth in SEQ ID NO: 2. In another embodiment, ADAMTS-5 is a human ADAMTS-5 protein which comprises amino acid 1 through amino acid 518 of the sequence set forth in SEQ ID NO: 4. In one embodiment, mature human ADAMTS-6 protein comprises amino acid 245 through amino acid 860 of SEQ ID NO: 6. In one embodiment, mature human ADAMTS-7 protein comprises amino acid 233 through amino acid 997 of the sequence set forth in SEQ ID NO: 8. In one embodiment, mature ADAMTS-8 protein is a mouse protein which comprises amino acid 229 through amino acid 905 of the sequence set forth in SEQ ID NO: 10. In another embodiment, ADAMTS-8 protein is a human protein which comprises amino acid 1 through amino acid 245 of the sequence set forth in SEQ ID NO: 12. In one embodiment, mature ADAMTS-9 protein

is a human protein which comprises amino acid 236 through amino acid 1882 of the sequence set forth in SEQ ID NO: 14. In another embodiment, ADAMTS-9 protein is a mouse protein which comprises amino acid 1 through amino acid 974 of the sequence set forth in SEQ ID NO: 16. In one embodiment, mature ADAMTS 10 protein is a human protein which comprises amino acid 212 through amino acid 1081 of the sequence set forth in SEQ ID NO: 18. In another embodiment, ADAMTS-10 protein is a mouse protein which comprises amino acid 1 through amino acid 547 of the sequence set forth in SEQ ID NO: 20.

The present invention also provides isolated polynucleotides which encode an ADAMTS-N protein or a variant thereof, polynucleotide sequences complementary to such polynucleotides, vectors containing such polynucleotides, and host cells transformed or transfected with such vectors. The present invention also relates to antibodies which are immunospecific for one or more of the ADAMTS-N proteins. The present invention also relates to a protein referred to hereinafter as ADAMTS-R1 (ADAM-T-S Related protein-1) and the polynucleotides which encode such protein. In one embodiment, the ADAMTS-R1 protein comprises amino acid 1 through amino acid 525 of the sequence set forth in SEQ. ID NO: 22.

Brief Description of the Drawings

Figure 1 shows an amino acid sequence (SEQ ID NO:2) of a full-length mouse ADAMTS-5 protein and a nucleic acid sequence (SEQ ID NO: 1) which encodes such protein.

Figure 2 shows an amino acid sequence (SEQ ID NO:4) of a partial human ADAMTS-5 protein and a nucleic acid sequence (SEQ ID NO: 3) which encodes such protein.

Figure 3 shows an amino acid sequence (SEQ ID NO:6) of a full-length human ADAMTS-6 protein and a nucleic acid sequence (SEQ ID NO:5) which encodes such protein.

Figure 4 shows an amino acid sequence (SEQ ID NO:8) of a full-length human ADAMTS-7 protein and a nucleic acid sequence (SEQ ID NO:7) which encodes such protein.

Figure 5 shows an amino acid sequence (SEQ ID NO: 10) of a full-length mouse ADAMTS-8 protein and a nucleic acid sequence (SEQ ID NO:9) which encodes such protein.

Figure 6 shows an amino acid sequence (SEQ ID NO: 12) of a partial human ADAMTS-8 protein and a nucleic acid sequence (SEQ ID NO: 11) which encodes such amino acid sequence.

Figure 7 shows an amino acid sequence (SEQ ID NO: 14), of a full-length human ADAMTS-9 protein and a nucleic acid sequence (SEQ ID NO: 13) Which encodes such protein.

Figure 8 shows an amino acid sequence (SEQ ID NO: 16) of a partial mouse ADAMTS-9 protein and a nucleic acid sequence (SEQ ID NO: 15) which encodes such amino acid sequence.

Figure 9 shows an amino acid sequence (SEQ ID NO:18) of a full-length human ADAMTS-10 protein and a nucleic acid sequence (SEQ ID NO: 17) which encodes such protein.

Figure 10 show's an amino acid sequence (SEQ ID NO:20) of a partial mouse ADAMTS-10 protein and a nucleic acid sequence (SEQ ID NO: 19) which encodes such amino acid sequence.

Figure 11 shows an amino acid sequence (SEQ ID NO:22) of a full length ADAMTS-R1 protein and a nucleic acid sequence (SEQ ID NO: 21) which encodes such protein.

Figure 12 depicts the cloning strategy used for isolation of a. mouse and human ADAMTS-5 cDNAs b. human ADAMTS-6 cDNA and c. human ADAMTS-7 cDNA. The domain organization of each protein relative to the cDNA clones (thin line) is shown as is the extent of overlap between clones. The original I.M.A.G.E. clones are underlined. Intronic regions of incompletely spliced transcripts are shown by the angl d

dotted lines. DNA scale marker (in bp) and amino acid scale marker are at upper right. Location of the probe used for *in situ* hybridization (ISH) is shown in a.

Figure 13 shows the predicted amino acid sequences of a. the mouse and human ADAMTS-5 proteins (alignment shows mouse sequence above, partial human sequence below) b. ADAMTS-6, and c. ADAMTS-7. The active-site sequences and proposed Met-turn are enclosed in boxes. Potential furin cleavage site(s) are indicated by arrows. Thrombospondin type-1 modules are underlined. Potential sites for N-inked glycosylation are overlined. Cysteine residues within the context of an MMP-like "cysteine switch" are indicated by the solid circles. Other cysteine residues are indicated by asterisks. The prodomain extends until the furin cleavage site, and the catalytic domain extends from the furin cleavage site to the approximate start of the disintegrin-like sequence (Dis). The start of the spacer domain is indicated; the region between the N-terminal TS domain and the spacer domain is the cysteine-rich domain. The single letter amino acid code is used.

Figure 14 shows Northern analysis of expression of ADAMTS-5, 6 and 7. RNA kilobase markers are shown at left of each autoradiogram, and tissue origin is indicated above each lane. a. Mouse embryo northern blots. b. Human multiple adult tissue northern blots.

Figure 15 is a schematic representation of the domain structure of ADAMTS-R1 protein as compared to ADAMTS-1 protein.

Figure 16 shows an amino acid sequence (SEQ ID NO: 24) of an alternative embodiment of a full-length human ADAMTS-10 protein and a nucleic acid sequence (SEQ ID NO: 23) which encodes such protein.

Figure 17 shows an amino acid sequence (SEQ ID NO: 26) of an alternative embodiment of human ADAMTS-9, which is a full-length protein designated as human ADAMTS-9b and a nucleic acid sequence

(SEQ ID NO: 25) which encodes such protein.

Figure 18 is a schematic representation of the domain structure of human ADAMTS-9b protein as compared to human and mouse ADAMTS-9 protein.

Detailed Description of the Invention

ADAMTS-N Proteins

The present invention relates to novel, isolated, substantially purified, mammalian proteins belonging to the ADAMTS subfamily of metalloproteases. As used herein, the term "substantially purified" refers to a protein that is removed from its natural environment, isolated or separated, and at least 60% free, preferably 75% free, and most preferably 90% free from other components with which it is naturally associated.

The novel mammalian proteins are ADAMTS-5, ADAMTS-6, ADAMTS-7, ADAMTS-8, ADAMTS-9 and ADAMTS-10, collectively ADAMTS-N. In one embodiment, the ADAMTS-5 protein is a mature mouse protein which comprises amino acid 231 through amino acid 930 of the sequence set forth in SEQ ID NO: 2. In another embodiment, the ADAMTS-5 protein is a human protein which comprises amino acid 1 through amino acid 518 of the sequence set forth in SEQ ID NO: 4. In one embodiment, ADAMTS-6 protein is a mat-Lire human protein which comprises amino acid 245 through amino acid 860 of SEQ ID NO:6. In one embodiment, the ADAMTS-7 protein is a mature human protein which comprises amino acid 233 through amino acid 997 of the sequence set forth in SEQ ID NO: 8. In one embodiment, the ADAMTS-8 protein is a mature mouse protein which comprises amino acid 229 through amino acid 905 of the sequence set forth in SEQ ID NO: 10. In another embodiment, the ADAMTS-8 protein is a human protein which comprises amino acid 1 through amino acid 245 of the sequence set forth in SEQ ID NO: 12. In one embodiment, the ADAMTS-9 is a mature human protein which comprises amino acid 236 through amino acid 1882 of the sequ nce set

forth in SEQ ID NO: 14. In another embodiment, the ADAMTS-9 protein is a mouse protein which comprises amino acid 1 through amino acid 874 of the sequence set forth in SEQ ID NO: 16. In another embodiment, the ADAMTS-9 designated ADAMTS-9b is a human protein which is comprised of 1934 amino acids as set forth in SEQ ID NO 26. In one embodiment, the ADAMTS-10 protein is a mature human protein which comprises amino acid 212 through amino acid 1081 of the sequence set forth in SEQ ID NO: 18. In another embodiment the ADAMTS- 10 protein is a mouse protein which comprises amino acid 1 through amino acid 525 of the sequence set forth in SEQ ID NO:20. In another embodiment, the ADAMTS-10 protein is a human protein which is comprised of 1072 amino acids as set forth in SEQ ID NO 24.

All of the novel ADAMTS-N proteins starting at the amino terminus comprise a signal sequence followed by a putative pro region which contains a consensus sequence for furin cleavage (except for ADAMTS-10), a catalytic domain, a domain of 60-90 residues with 35 to 45% similarity to snake venom disintegrins, a TS module, a cysteine rich domain containing multiple conserved cysteine residues, a spacer domain, and one or multiple C terminal TS modules. (See Figure 12.) As determined using the BLAST software from the National Center for Biotechnology Information, the predicted mature forms of the ADAMTS-N proteins show an overall 20-30% similarity to each other and to ADAMTS-1-4, although this may be considerably higher or lower for individual domains as described below.

The ADAMTS-N proteins also encompass variants of the ADAMTS-N proteins shown in Figs. 1-10. A "variant" as used herein, refers to a protein whose amino acid sequence is similar to one of the amino acid sequences shown in Figs. 1-10, hereinafter referred to as the reference amino acid sequence, but does not have 100% identity with the ref rence sequence. Th variant protein has an alter d sequence

in which one or more of the amino acids in the reference sequence is deleted or substituted, or one or more amino acids are inserted into the sequence of the reference amino acid sequence. As a result of the alterations, the variant protein has an amino acid sequence which is at least 95% identical to the reference sequence, preferably, at least 97% identical, more preferably at least 98% identical, most preferably at least 99% identical to the reference sequence. Variant sequences which are at least 95% identical have no more than 5 alterations, i.e. any combination of deletions, insertions or substitutions, per 100 amino acids of the reference sequence. Percent identity is determined by comparing the amino acid sequence of the variant with the reference sequence using MEGALIGN project in the DNA STAR program. Sequences are aligned for identity calculations using the method of the software basic local alignment search tool in the BLAST network service (the National Center for Biotechnology Information, Bethesda, MD) which employs the method of Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990) *J. Mol. Biol.* 215, 403-410. Identities are calculated by the Align program (DNAstar, Inc.) In all cases, internal gaps and amino acid insertions in the candidate sequence as aligned are not ignored when making the identity calculation.

While it is possible to have nonconservative amino acid substitutions, it is preferred that the substitutions be conservative amino acid substitutions, in which the substituted amino acid has similar structural or chemical properties with the corresponding amino acid in the reference sequence. By way of example, conservative amino acid substitutions involve substitution of one aliphatic or hydrophobic amino acids, e.g. alanine, valine, leucine and isol ucine, with another; substitution of on hydroxyl-containing amino acid, e.g. serine and threonine, with another; substitution of

one acidic residue, e.g. glutamic acid or aspartic acid, with another; replacement of one amide-containing residue, e.g. asparagine and glutamine, with another; replacement of one aromatic, residue, e.g. phenylalanine and tyrosine, with another; replacement of one basic residue, e.g. lysine, arginine and histidine, with another; and replacement of one small amino acid, e.g., alanine, serine, threonine, methionine, and glycine, with another.

The alterations are designed not to abolish the immunoreactivity of the variant protein with antibodies that bind to the reference protein. Guidance in determining which amino acid residues may be substituted, inserted or deleted without abolishing immunoreactivity of the variant protein with an antibody specific for the respective reference protein are found using computer programs well known in the art, for example, DNASTAR software.

The ADAMTS-N proteins also encompass fusion proteins comprising an ADAMTS-N protein and a tag, i.e., a second protein or one or more amino acids, preferably from about 2 to 65 amino acids, more preferably from about 34 to about 62 amino acids, which are added to the amino terminus of, the carboxy terminus of, or any point within the amino acid sequence of an ADAMTS-N protein, or a variant of such protein. Typically, such additions are made to stabilize the resulting fusion protein or to simplify purification of an expressed recombinant form of the corresponding ADAMTS-N protein or variant of such protein. Such tags are known in the art. Representative examples of such tags include sequences which encode a series of histidine residues, the epitope tag FLAG, the Herpes simplex glycoprotein D, beta-galactosidase, maltose binding protein, or glutathione S-transferase.

The ADAMTS-N proteins also encompass ADAMTS-N proteins in which one or more amino acids, preferably no more than 10 amino acids, in

the respective ADAMTS-N protein are altered by posttranslation processes or synthetic methods. Examples of such modifications include, but are not limited to, acetylation, amidation, ADP-ribosylation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or a lipid, cross-linking gamma-carboxylation, glycosylation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, sulfation, and transfer-RNA mediated additions of amino acids to proteins such as arginylation and ubiquitination.

The ADAMTS-N proteins are immunogenic and, thus, are useful for preparing antibodies. Such antibodies are useful for identifying and diagnosing disorders which are associated with decreased expression or activity or increased expression of an ADAMTS-N protein. The ADAMTS-N protein may also be useful for treating such disorder.

Diseases involving enhanced or depressed proteolyisis of the core proteins of the extracellular may involve enhanced expression or activity or decreased expression or activity of one or more ADAMTS-N proteins. Thus, ADAMTS-N proteins may be used to identify drugs, polypeptides, auto-antibodies, or other natural compounds which bind to an ADAMTS-N protein with sufficient affinity to block or facilitate its activity. The activity of the ADAMTS-N protein is assayed in the presence and the absence of the putative inhibitor or facilitator using any of a variety of protease assays known in the art. In general, the activity of the ADAMTS-N protein is assayed through the use of a peptide or protein substrate having a known or putative cleavage site for the ADAMTS-N protein. To detect cleavage or to monitor the extent of cleavage, the substrate is tagged in a manner which provides a detectable signal upon cleavage. For example, the substrat may be tagged with a fluorescent group on one

side of the cleavage site and with a fluorescence-quenching group on the opposite side of the cleavage site. Upon cleavage by the substrate, quenching is eliminated and a detectable signal is produced. Alternatively, the substrate is tagged with a colorimetric leaving group that more strongly absorbs upon cleavage. Agents which block ADAMTS-N-catalyzed cleavage of a protein substrate may be administered to a subject to block proteolysis of the corresponding protein substrate.

ADAMTS-R1 Protein

The present invention also relates to a protein, referred to hereinafter as "ADAMTS-R1". From its amino to its carboxyl terminus, ADAMTS-R1 comprises a signal peptide sequence, a TS1 module, a cysteine-rich domain, a spacer domain, and three TS1 modules. Thus, ADAMTS-R1 has a structure which is related to or similar to an ADAMTS-N protein, but which lacks a catalytic domain and a disintegrin-like domain. In one embodiment, ADAMTS-R1, protein comprises amino acid 1 through amino acid 525 of the amino acid sequence, SEQ ID NO:22, shown in Fig. 11. Such protein has a 30-40% overall sequence identity with similar regions of the ADAMTS-N proteins. The ADAMTS-R1 proteins also encompass variants of the amino acid sequence shown in Fig. 11 and fusion proteins which contain the amino acid sequence shown in Fig. 11 or a variant thereof. On the basis of its domain organization, it is expected that ADAMTS-R1 can bind to extracellular matrix or cell surface molecules, including ADAMTS-N substrates. Thus, it is expected that ADAMTS-R1 can be used as an cell-matrix or cell-cell adhesion molecule or an ADAMTS-N competitive inhibitor. The ADAMTS-R1 proteins are also useful for preparing antibodies. Such antibodies are useful for identifying tissues where ADAMTS-R1 is expressed and for diagnosing disorders which are associated with decreased

expression or increased expression of. an ADAMTS-R1 protein.

Polynucleotides

The present invention also provides isolated polynucleotides which encode the mammalian ADAMTS-N proteins and the mammalian ADAMTS-R1 protein. Figure 1 shows one embodiment of a polynucleotide, SEQ ID NO: 1, which encodes the full-length mouse ADAMTS-5 protein. Figure 2 shows one embodiment of a polynucleotide; SEQ ID NO: 3, which encodes a partial human ADAMTS-5 protein. Figure 3 shows one embodiment of a polynucleotide; SEQ ID NO: 5, which encodes a full-length human ADAMTS-6 protein. Figure 4 shows one embodiment of a polynucleotide; SEQ ID NO: 7, which encodes a full-length human ADAMTS-7 protein. Figure 5 shows one embodiment of a polynucleotide; SEQ ID NO: 9, which encodes a full-length mouse ADAMTS-8 protein. Figure 6 shows one embodiment of a polynucleotide; SEQ ID NO: 11, which encodes a partial human ADAMTS-8 protein. Figure 7 shows one embodiment of a polynucleotide; SEQ ID NO: 13, which encodes a full-length human ADAMTS-9 protein. Figure 8 shows one embodiment of a polynucleotide; SEQ ID NO: 15, which encodes a partial ADAMTS-9 protein. Figure 9 shows one embodiment of a polynucleotide; SEQ ID NO: 17, which encodes a full-length human ADAMTS-10 protein. Figure 10 shows one embodiment of a polynucleotide; SEQ ID NO: 19, which encodes a partial mouse ADAMTS-10 protein. Figure 11 shows one embodiment of a polynucleotide; SEQ ID NO: 21, which encodes a full-length ADAMTS-R1 protein.

Due to the known degeneracy of the genetic code wherein more than one codon can encode the same amino acid, a DNA sequence may vary from that shown in SEQ ID NO: 1 and still encode an ADAMTS-5 protein having the amino acid sequence of SEQ ID NO: 2. Similarly, a DNA sequence may vary from that shown in SEQ ID NO:5, and still encode an ADAMTS-6 protein having th amino acid sequence set forth

in SEQ ID NO:6. Similarly a DNA sequence may vary from that shown in SEQ ID NOS: 7, 9, 11, and 13, and still encode the amino acid sequences shown in SEQ ID NOS: 8, 10, 12, and 14, respectively. Such variant DNA sequence may result from silent mutations, such as for example those that occur during PCR amplification or from deliberate mutagenesis of a native sequence.

The present polynucleotides also encompass polynucleotides having sequences that are capable of hybridizing to the nucleotide sequences of FIGS 1 - 11 under stringent conditions, preferably highly stringent conditions. Hybridization conditions are based on the melting temperature™ of the nucleic acid binding complex or probe, as described in Berger and Kimmel (1987) Guide to Molecular Cloning Techniques, Methods in Enzymology, vol 152, Academic Press. The term "stringent conditions, as used herein, is the "stringency" which occurs within a range from about Tm-5 (5° below the melting temperature of the probe) to about 20° C below Tm. As used herein "highly stringent" conditions employ at least 0.2 x SSC buffer and at least 65° C. As recognized in the art, stringency conditions can be attained by varying a number of factors such as the length and nature, i.e., DNA or RNA, of the probe; the length and nature of the target sequence, the concentration of the salts and other components, such as formamide, dextran sulfate, and polyethylene glycol, of the hybridization solution. All of these factors may be varied to generate conditions of stringency which are equivalent to the conditions listed above.

The present polynucleotides also encompasses alleles of the ADAMTS-N and ADAMTS-R1 encoding sequences. As used herein, an allele or allelic sequence is an alternative form of an ADAMTS-N or ADAMTS-R1 encoding sequence which is present at the same gene locus. The allele may r sult from one r more mutations in the ADAMTS-N or

ADAMTS-R1 encoding sequence. Such mutations typically arise from natural addition, deletion of substitution of nucleotides in the open reading frame sequences. Any gene which encodes an ADAMTS-N protein or ADAMTS-RI protein may have none, one, or several allelic forms. Such alleles are identified using conventional techniques, such as for example screening libraries with probes having sequences identical to or complementary with one or more ADAMTS-N polynucleotides.

The present polynucleotides also encompass altered polynucleotides which encode ADAMTS-N proteins, ADAMTS-R1 proteins, and variants thereof. Such alterations include deletions, additions, or substitutions. Such alterations may produce a silent change and result in an ADAMTS-N protein having the same amino acid sequence as the ADAMTS-N protein encoded by the unaltered polynucleotide. Such alterations may produce a nucleotide sequence possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eucaryotic host may be incorporated into the nucleotide sequences showing Figures 1 -11 to increase the rate of expression of the proteins encoded by such sequences. Such alterations may also introduce new restriction sites into the sequence or result in the production of an ADAMTS-N or ADAMTS-RI variant. Typically, such alterations are accomplished using site-directed mutagenesis.

The polynucleotides are useful for producing ADAMTS-N or ADAMTS-R1 proteins. For example, an RNA molecule encoding an ADAMTS-N protein is used in a cell-free translation systems to prepare such protein. Alternatively, a DNA molecule encoding an ADAMTS-N protein is introduced into an expression vector and used to transform cells. Suitable expression vectors include for example chromosomal, nonchromosomal and synthetic DNA sequences, e.g., derivatives of

SV40, bacterial plasmids, phage DNAs; yeast plasmids, vectors derived from combinations of plasmids and phage DNAs, viral DNA such as vaccinia, adenovirus, fowl pox virus, pseudorabies, baculovirus, and retrovirus. The DNA sequence is introduced into the expression vector by 5 conventional procedures.

Accordingly, the present invention also relates to recombinant constructs comprising one or more of the present polynucleotide sequences. Suitable constructs include, for example, vectors, such as a plasmid, phagemid, or viral vector, into which a sequence that encodes an ADAMTS-N protein or an ADAMTS-R1 protein has been inserted. In the expression vector, the DNA sequence which encodes the ADAMTS-N protein is operatively linked to an expression control sequence, i.e., a promoter, which directs mRNA synthesis. Representative examples of such promoters, include the LTR or SV40 promoter, the *E. coli* lac or trp, the phage lambda PL promoter and other promoters known to control expression of genes in prokaryotic or eukaryotic cells or in viruses. The promoter may also be the natural promoter of the ADAMTS-N encoding sequence. The expression vector, preferably, also contains a ribosome binding site for translation initiation and a transcription terminator. Preferably, the recombinant expression vectors also include an origin of replication and a selectable marker, such as for example, the ampicillin resistance gene of *E. coli* to permit selection of transformed cells, i.e. cells that are expressing the heterologous DNA sequences. The polynucleotide sequence encoding the ADAMTS-N protein is incorporated into the vector in frame with translation initiation and termination sequences.

The polynucleotides encoding an ADAMTS-N or ADAMTS-R1 protein are used to express recombinant protein using techniques well known in th art. Such techniques are described in Sambrook, J. et al

(1989) Molecular Cloning A Laboratory Manual, Cold Spring Harbor Press, Plainview, N.Y. and Ausubel, F. M. et al. (1989) Cuurent Protocols in Molecular Biology, John Wile & Sons, New York, NY.

Polynucleotides encoding an ADAMTS-N or ADAMTS-R1 protein may also be used for diagnostic purposes. The polynucleotides may be used to detect and quantify ADAMTS-N or ADAMTS-R1 gene transcripts in biopsied tissues in which enhanced expression or reduced expression of the corresponding ADAMTS-N or ADAMTS-RI gene is correlated with a disease. The diagnostic assay may be used to determine whether expression is absent, present, or altered and to determine whether certain therapeutic agents modulate expression of the corresponding ADAMTS-N or ADAMTS-R1 gene.

Also encompassed by the present invention, are single stranded polynucleotides, hereinafter referred to as antisense polynucleotides, having sequences which are complementary to the DNA and RNA sequences which encode the ADAMTS-N or ADAMTS-R1 proteins. The term complementary as used herein refers to the natural binding of the polynucleotides under permissive salt and 5 temperature conditions by base pairing.

The present invention also encompasses oligonucleotides that are used as primers in polyrnerase chain reaction (PCR) technologies to amplify transcripts of the genes which encode the ADAMTS-N and ADAMTSR-1 proteins or portions of such transcripts. Preferably, the primers comprise 18-30 nucleotides, more preferably 19-25 nucleotides. Preferably, the primers have a G+C content of 40% or greater. Such oligonucleotides are at least 98% complementary with a portion of the DNA strand, i.e., the sense strand, which encodes the respective ADAM-TS family protein or a portion of its corresponding antisense strand. Preferably, the primer has at least 99% complementarity, more preferably 100% complementarity, with such

sense strand or its corresponding antisense strand. Primers which are which have 100% complementarity with the antisense strand of a double-stranded DNA molecule which encodes an ADAMTS-N protein have a sequence which is identical to a sequence contained within the sense strand. The identity of primers which are 15 nucleotides in length and have full complementarity with a portion of the antisense strand of a double-stranded DNA molecule which encodes the ADAMTS-N protein is determined using the nucleotide sequences, shown in FIG I - 11 and described by the general formula a-b; where a is any integer between I and the position number of the nucleotide which is located 15 residues upstream of the 3' end of the sense or antisense strand of the cDNA sequences shown in FIG 1 -11; where b is equal to a+14; and where both a and b correspond to the positions of nucleotide residues of the cDNA sequences shown in FIGS 1 - 11.

The present invention also encompasses oligonucleotides that are useful as hybridization probes for for isolating and identifying cDNA clones and genomic clones encoding the ADAMTS-N or ADAMTS-R1 protein or allelic forms thereof. Such hybridization probes are also useful for detecting transcripts of the genes which encode the ADAMTS-N family proteins or for mapping of the genes which encode the ADAMTS-N proteins Preferably, such oligonucleotides comprise at least 210 nucleotides, more preferably at least 230, most preferably from about 210 to 280 nucleotides. Such hybridization probes have a sequence which is at least 90% complementary with a sequence contained within the sense strand of a DNA molecule which encodes an ADAMTS-N protein or ADAMTS-R1 protein or with a sequence contained within its corresponding antisense strand. Such hybridization probes bind to the sense strand under stringent conditions. The term "stringent conditions" as used herein is the binding which occurs within a range from about Tin 5'C (5'C below the melting temperature

Tm of the probe) to about 20°C to 25°C below Tm. The probes are used in Northern assays to detect transcripts of ADAMTS-N homologous genes and in Southern assays to detect ADAMTS-N homologous genes. The identity of probes which are 200 nucleotides 5 in length and have full complementarity with a portion of the antisense strand of a double-stranded DNA molecule which encodes the ADAMTS-N protein is determined using the nucleotide sequences shown in FIG 1 - 10 and described by the general formula a-b; where a is any integer between I and the position number of the nucleotide which is located 200 residues upstream of the 3' end of the sense or antisense strand of the cDNA sequences shown in FIG 1 -10; b is equal to a +200; and where both a and b correspond to the positions of nucleotide residues of the cDNA sequences shown in FIG 1-10.

Such probes or primers are also useful for identifying tissues or cells in which the corresponding ADAMTS-N or ADAMTS-R1 gene is preferentially expressed either constitutively or at particular state of tissue differentiation or development or in disease states. Expression of the ADAMTS-N or ADAMTS-R1 gene in a particular tissue or group of cells is determined using conventional procedures including, but not limited to, Northern analysis, in situ hybridization to RNA or RT-PCR amplification. Isolated polynucleotides encoding an ADAMTS-N or ADAMTS-R1 protein are also useful as chromosome markers to map linked gene positions, to identify chromosomal aberrations such as translocations, inversions and trisomies, to compare with endogenous DNA sequences in patients to identify potential genetic disorders, and as probes to hybridize and thus discover novel, related DNA sequences. For use in such studies and assays, the probes may be labeled with radioisotopes, fluorescent labels, or enzymatic labels. The assays includ , but are not limited to, Southern blot, in situ hybridization to DNA in cells

and chromosomes, PCR, and allele specific hybridization.

Antibodies

In another aspect, the present invention relates to antibodies which are specific for and bind to the ADAMTS-5 protein, the ADAMTS-6 protein, the ADAMTS-7 protein, the ADAMTS-8 protein, the ADAMTS-9 protein, the ADAMTS-10 protein, or the ADAMTS-R1 protein. Such antibodies are useful research tools for identifying *tissues that contain elevated levels of the respective protein and for purifying the respective protein from cell or tissue extracts, medium of cultured cells, or partially purified preparations of intracellular and extracellular proteins by affinity chromatography. Such antibodies are also useful for identifying and diagnosing diseases associated with elevated or reduced levels of an ADAMTS-N protein or ADAMTS-R1 protein. Such antibodies are also useful for monitoring the effect of therapeutic agents on the synthesis and secretion of ADAMTS-N proteins by cells in vitro and in vivo. Such antibodies may also be employed in procedures, such as co-immunoprecipitation and co-affinity chromatography, for identifying other proteins, activators and inhibitors which bind to an ADAMTS-N or ADAMTS-R1 protein.

The present invention also provides a method for detecting an ADAMTS-N or ADAMTS-R1 protein, in a bodily sample from a patient using antibodies immunospecific for an ADAMTS-N or ADAMTS-R1 protein. The method comprises contacting the antibody with a sample taken from the patient; and assaying for the formation of a complex between the antibody and the corresponding ADAMTS-N or ADAMTS-R1 protein present in the sample. The sample may be a tissue or a biological fluid, including but not limited to whole blood, serum, synovial fluid, stool, urine, cerebrospinal fluid, semen, diagnostic wash s from trachea, stomach and other bowel segments, tissue biopsies or excised

tissue, cells obtained from swabs and smears. To monitor changes in expression of the ADAMTS-N protein during fetal development and pregnancy, it is preferred that the sample be amniotic fluid. To monitor changes in expression of the ADAMTS-N protein during joint disorders, the preferred sample is synovial fluid. To monitor changes in expression of ADAMTS-N proteins during cancer, the preferred samples include, but are not limited to, serum, body fluids, or biopsy tissue. To monitor changes in expression of ADAMTS-N proteins during inflammation the preferred samples include, but are not limited to, serum, body fluids, or biopsy tissue.

The sample may be untreated, or subjected to precipitation; fractionation, separation, or purification before combining with the anti-ADAMTS-N protein antibody. For ease of detection, it is

preferred that isolated proteins from the sample be attached to a substrate such as. a column, plastic dish, matrix, or membrane, preferably nitrocellulose. Preferably, the detection method employs an enzyme-linked immunosorbent assay (ELISA) or a Western immunoblot procedure.

Interactions between an ADAMTS-N protein in the sample and the corresponding anti ADAMTS-N antibody are detected by radiometric, colorimetric, or fluorometric means, size separation, or precipitation. Preferably, detection of the antibody-ADAMTS-N protein complex is by addition of a secondary antibody that is coupled to a detectable tag, such as for example, an enzyme, fluorophore, or chromophore. Formation of the complex is indicative of the presence of the ADAMTS-N protein i n the test sample. Thus, the method is used to determine whether there is a decrease or increase in the levels of the ADAMTS-N protein in a test sample as compared to levels of the ADAMTS-N protein in a control sample and to quantify the amount of the ADAMTS-N protein in the test sample.

Deviation between control and test values establishes the parameters for diagnosing the disease.

Preparing the ADAMTS-N proteins and the ADAMTS-R1 protein

The ADAMTS-N proteins and the ADAMT-SR1 protein may be produced by conventional peptide synthesizers. The ADAMTS-N proteins and the ADAMTS-R1 protein may also be produced using cell-free translationsystems and RNA molecules derived from DNA constructs that encode an ADAMTS-N protein or an ADAMTS-RI protein. Alternatively, ADAMTS-N proteins are made by transfecting host cells with expression vectors that comprise a DNA sequence that encodes the respective ADAMTS-N protein and then inducing expression of the protein in the host. cells. For recombinant production, recombinant constructs comprising one or more of the sequences which encode the ADAMTS-N protein or a variant thereof are introduced into host cells by conventional methods such as calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape lading, ballistic introduction or infection.

The ADAMTS-N protein and the ADAMTS-R1 protein may be expressed in suitable host cells, such as for example, mammalian cells, yeast, bacteria, insect cells or other cells under the control of appropriate promoters using conventional techniques. Suitable hosts include, but are not limited to, E. coli, P. pastoris, Cos cells and 293 HEK cells. Following transformation of the suitable host strain and growth of the host strain to an appropriate cell density, the cells are harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification of the ADAMTS-N protein or the ADAMTS-R1 protein.

Conventional procedures for isolating recombinant proteins from transformed host cells, such as isolation by initial extraction from

cell pellets or from cell culture medium, followed by salting-out, and one or more chromatography steps, including aqueous ion exchange chromatography, size exclusion chromatography steps, and high performance liquid chromatography (HPLC), and affinity chromatography may be used to isolate the recombinant ADAMTS-N protein or ADAMTS R1 protein

Preparation of Antibodies

The ADAMTS-N proteins, and variants thereof are used as immunogens to produce antibodies immunospecific for one or more ADAMTS-N protein. The term "immunospecific" means the antibodies have substantially greater affinity for one or more ADAMTS-N protein than for other proteins. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, and Fab fragments.

Antibodies are also prepared using an oligopeptide having a sequence which is identical to a portion of the amino acid sequence of an ADAMTS-N protein. Preferably the oligopeptide has an amino acid sequence of at least five amino acids, and more preferably, at least 10 amino acids that are identical to a portion of the amino acid sequence of an ADAMTS-N protein. Such peptides are conventionally fused with those of another protein such as keyhole limpet hemocyanin and antibody produced against the chimeric molecule. One preferred oligopeptide for preparing an antibody to mouse ADAMTS-5 has the sequence (C)HIKVRQFKAKDQTRF, SEQ ID NO: 30. Another preferred oligopeptide for preparing an antibody to ADAMTS-5 is CEAKNGYQSDAKGVKTFVEWVPKYAG, SEQ ID NO: 3 1. One preferred oligopeptide for preparing an antibody to ADAMTS-6 has the sequence SVSIERFVETLVVADK(C), SEQ ID NO:23. One preferred oligopeptide for preparing an antibody to ADAMTS-7 has the sequence (C)EVAEAANFLALRSEDPEKY, SEQ ID NO:24. One preferr d oligopeptide for

preparing an antibody to ADAMTS-8 has the sequence CVKEDVENPKAVVDGDWGP, SEQ ID NO:25. One preferred oligopeptide for preparing an antibody to ADAMTS-9 has the sequence QHPFQNEDYRPRSASPSRTH, SEQ ID NO:26. Another preferred oligopeptide for preparing an antibody to ADAMTS-9 has the sequence PQNCKEVKRLKGASEDGEYF, SEQ ID NO:27. One preferred oligopeptide for preparing an antibody for ADAMTS-R1 has the sequence QELEEGAAVSEEPS, SEQ ID NO:28. Another preferred oligopeptide for preparing an antibody for ADAMTS-R1 has the sequence YYPENIKPKPKLQE; SEQ ID NO:29.

Polyclonal antibodies are generated using conventional techniques by administering the ADAMTS-N protein or achimeric molecule to a host animal. Depending on the host species, various adjuvants may be used to increase immunological response. Among adjuvants used in humans, Bacilli Calmette-Guerin (BCG), and Corynebacterium parvum. are especially preferable. Conventional protocols are also used to collect blood from the immunized animals and to isolate the serum and or the IgG fraction from the blood.

For preparation of monoclonal antibodies, conventional hybridoma techniques are used. Such antibodies are produced by continuous cell lines in culture. Suitable techniques for preparing monoclonal antibodies include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV hybridoma technique.

Various immunoassays may be used for screening to identify antibodies having the desired specificity. These include protocols which. involve competitive binding or immunoradiometric assays and typically involve the measurement of complex formation between the respective ADAMTS-N protein and the antibody.

Polynucleotides that encode ADAMTS-N proteins

Polynucleotides comprising sequences encoding an ADAMTS-N

protein or an ADAMTS-R1 protein may be synthesized in whole or in part using chemical methods. Polynucleotides which encode an ADAMTS-N protein, particularly alleles of the genes which encode the ADAMTS-N protein, may be obtained by screening a genomic library or cDNA library with a probe comprising sequences identical or complementary to the sequences shown in Figures 1 - 10 or with antibodies immunospecific for a ADAMTS-N protein to identify clones containing such polynucleotide.

Example 1 ADAMTS-512 protein

A cDNA encoding mouse ADAMTS-5 protein was obtained using IMAGE Clone 569515, purchased from Research Genetics, Huntsville, Alabama and 7 day old mouse embryo cDNA library from Clontech, Palo Alto, CA. A cDNA encoding human ADAMTS-5 protein was obtained using IMAGE Clone 345484 purchased from Research Genetics, Huntsville, Alabama and a human fetal brain cDNA from Clontech. The clone inserts were sequenced in their entirety. Using oligonucleotide primers based on the sequences at the ends of the. clone inserts as template, successive rounds of RACE (Rapid Amplification of cDNA Ends) by PCR was performed at 5' and 3 ends. RACE primers were generated 50-200 bp from the ends of the sequences so that the contiguity of RACE clones with the I.M.A.G.E. clone could be clearly established. A single round of 5' and 3' 20 RACE sufficed for cloning of the entire coding sequence of the mouse ADAMTS-5 protein and part of the catalytic zinc binding site through to the stop codon of the human ADAMTS-5 protein. Primers were designed with calculated Tm>72°C and RACE was performed with nested primers for each amplification. PCR used the Advantage PCR reagents (Clontech, Palo Alto, CA); the polymerase mix contained both *Taq* polymerase as well as proofreading polymerase to minimize PCR errors and employed "hot-start" PCR for optimal efficiency. RACE used the following "touch-down" cycle

conditions; 95°C for 1 minute followed by 5 cycles of 95°C for 0.5 minutes, 72°C for 5 minutes, then 5 cycles of 95°C for 0.5 minutes, 70°C for 5 minutes and 20 cycles of 95°C for 0.5 minutes, 68°C for 5 minutes. The PCR products were analyzed by Southern blotting, initially using [α^{32}P]-dCTP labeled.

Hybridizing bands were ligated into pGEM-T Easy (Promega, Madison, WI) and individual clones were selected by another round of Southern analysis. Automated nucleotide sequencing of both strands of each clone were done at the Molecular Biotechnology Core of the Lerner Research Institute, Cleveland Clinic Foundation and nucleotide sequence data were analyzed using the DNAStar software. By integration of the overlapping sequences thus obtained, a contiguous nucleotide sequence was determined. The nucleotide sequence of the mouse ADAMTS-5 cDNA and the predicted amino acid sequence of the protein encoded by this cDNA are shown in Fig. 1. The nucleotide sequence of the human ADAMTS-5 cDNA and the predicted partial amino acid sequence of the protein encoded by this cDNA are shown in Fig. 2.

The predicted molecular mass (Mr) of the mature ADAMTS-5 protein is 73717.50 daltons. It is expected that the actual Mr of the active ADAMTS-5 protein is different due to post-translational modification, which could potentially increase the Mr. The predicted domain organization of ADAMTS-5 protein relative to the cloned cDNA is shown in Figure 12. The pro-domain of the full-length mouse ADAMTS-5 protein has 3 consensus cleavage signals for furin. The most carboxyl-terminal furin cleavage site in ADAMTS-5 predicts the processing site for generation of the mature protein The catalytic domain of the ADAMTS-5 protein contains eight cysteine residues and a reprolysin -zinc binding signature sequence, i.e., HEIGHLLGLSHD. Five cysteine residues are upstream of the zinc binding sequence,

while three residues are downstream, an arrangement that is shared with other ADAMTS members. The zinc binding signature is followed by a "Met-turn". The catalytic domain is followed by a domain with 35% similarity to snake venom disintegrins. The disintegrin domain contains eight cysteine residues. The first TS repeat contains 52 residues and is followed by a conserved cysteine-rich sequence termed the cysteine-rich domain, designated "CRD", to distinguish it from the cysteine-free spacer domain. The CRD contains ten conserved cysteines and demonstrates high sequence homology with the CRD of other ADAMTS-N proteins. The spacer domain of mouse ADAMTS-5 is 158 amino acids in length and is followed by a second TS module. ADAMTS-5 contains three potential glycosylation sites in the mature protease one of which is just upstream of the start of the spacer domain and the second lies within the spacer domain and the third is near the start of the disintegrin domain. The human ADAMTS-5 protein and the mouse ADAMTS-5 protein have 96% sequence identity. ADAMTS-5 bears 46% sequence identity to ADAMTS-4 (KIAA0688), which is characterized as being involved in catabolism of aggrecan core protein in arthritis and 60% identity to ADAMTS-1 which is involved in inflammation.

Example 2 ADAMTS-6

The nucleotide sequence of a human cDNA encoding the full-length ADAMTS-6 protein was obtained using IMAGE clone 742630, which encodes EST AA400393, and a human fetal brain cDNA from Clontech. RACE was performed as described above in Example 1. The I.M.A.G.E. clone 742630 contained an ORF flanked by consensus splice sequences, indicating the presence of introns. Two successive rounds of RACE at the 5' end and a single round of RACE at the 3' end provided the complete coding sequence of ADAMTS-6. The putative ATG codon is within a Kozak consensus sequence and encodes the first methionine within the ORF.

The nucleotide sequence of the ADAMTS-6 DNA is shown in Fig. 3 The predicted amino acid sequence, SEQ ID NO:6, of the ADAMTS-6 protein is also shown in Fig. 3. The predicted Mr of the full-length, unprocessed ADAMTS-6 protein is 97,115 daltons., and the predicted Mr of the mature ADAMTS-6 protein is 68412.10 daltons. The domain organization of the ADAMTS-6 protein is shown in Fig. 12. The pro-domain of the full-length ADAMTS-6 protein has one consensus cleavage signal for furin. The catalytic domain of the ADAMTS-6 contains six cysteine residues and the reprolysin -zinc binding signature sequence, HEIVHNFGMNHD, which is followed by a "Met-tum". The catalytic domain is followed by a domain with 35% similarity to snake venom disintegrins. The disintegrin domain contains eight cysteine residues. The first TS repeat contains 52 residues and is followed by a conserve CRD sequence which contains ten conserved cysteines and demonstrates high sequence homology with the CRD of other ADAMTS proteins. The spacer domain of ADAMTS-6 is 127 amino acids in length and is followed by a second TS module. ADAMTS-6 contains four potential glycosylation sites within the pyo-domain and two in the mature protease one of which is in the cysteine rich domain and the other of which is in the spacer domain. ADAMTS-6 bears 46% sequence identity to ADAMTS-1, which is involved in inflammation.

Example 3 ADAMTS-7.

The nucleotide sequence of a cDNA encoding an ADAMTS-7 protein was obtained using IMAGE clone 272098, which encodes EST N4.8032, and a human fetal brain cDNA from Clontech. RACE was performed as described above in Example 1. The I.M.A.G.E. clone 272098 encoded a putative pre-pro region and was extended in the 3'-direction by two successiv rounds of RACE. A typical signal peptide sequence lies downstream of th first methionine in the translated ORF. This

methionine codon lies within a satisfactory Kozak consensus for translation initiation.

The nucleotide sequence of the ADAMTS-7 cDNA is shown in Fig. 4. The predicted amino acid sequence, SEQ ID NO: 8, of the ADAMTS-7 protein is also shown in Fig. 4. The predicted Mr of the hill-length, unprocessed ADAMTS-7 protein is 116,607 daltons, and the predicted Mr of the mature ADAMTS-7 protein is 84005 daltons. The domain organization of the ADAMTS-7 protein is shown in Fig. 12. The pro-domain of the full-length ADAMTS-7 protein has one consensus cleavage signal for furin. The catalytic domain of the ADAMTS-7 protein contains eight cysteine residues and the reprolysin-zinc binding signature sequence, HELGHSFGIQHD, which is followed by a "Met-tum". The catalytic domain is followed by a domain with 30% similarity to snake venom disintegrins The disintegrin domain contains eight cysteine residues. The first TS repeat contains 52 residues and is followed by a conserved CRD sequence which contains ten conserved cysteines. The spacer domain of ADAMTS-7 is 221 amino acids in length and is followed by a second TS module and a short sequence containing two cysteine residues. ADAMTS-7 contains three potential glycosylation sites within the mature protease; one of which is just upstream of the spacer domain and one of which is within the spacer domain. ADAMTS-7 bears 35 % sequence identity to ADAMTS-1, which is characterized as being involved in inflammation and 32% identity to ADAMTS-2 which is a procollagen processing enzyme.

Example 4: ADAMTS-8

The nucleotide sequence of a cDNA encoding a full-length, mouse ADAMTS-8 protein was obtained using IMAGE clone 1260693, which encodes EST AA855532, and a mouse embryo cDNA from Clonetech. The nucleotide sequenc of a cDNA encoding a partial ADAMTS-8 human

protein was obtained using IMAGE clone 2119838, which encodes EST A1400905, and a human fetal brain cDNA library from Clontech. RACE was performed, as described above in Example 1. The nucleotide sequence of the cDNA encoding the full-length ADAMTS-8 mouse protein and the amino acid sequence of such protein is shown in Fig. 5. The nucleotide sequence of the cDNA encoding the partial ADAMTS-8 human protein and the amino acid sequence of such protein is shown in Fig. 6.

The predicted Mr of the full-length, unprocessed ADAMTS-8 mouse protein is 1260693 daltons, and the predicted Mr of the mature ADAMTS-8 protein is 68412.10 daltons. The pro domain of the full-length ADAMTS-8 protein has one consensus cleavage signal for furin. The catalytic domain contains eight cysteine residues and the reprolysm-zinc binding signature sequence, HELGHVLSMPHD, which is followed by a "Met-turn". The catalytic domain is followed by a domain with 20-30% similarity to snake venom disintegrins. The disintegrin-like domain contains eight cysteine residues. The first TS repeat is followed by a conserved CRD sequence which contains 10 conserved cysteines. The spacer domain of ADAMTS-8 is 146 amino acids in length and is followed by a second TS module. The ADAMTS-8 protein contains 4 potential glycosylation sites within the mature protease: one is in the cysteine-rich domain; one is in the catalytic domain; and two are in the disintegrin-like domain. ADAMTS-8 bears 46% sequence identity to ADAMTS-1 and 42% identity to ADAMTS-4.

<u>Example 5: ADAMTS-9</u>

The nucleotide sequence of a cDNA encoding a full-length, human ADAMTS-9 protein was obtained using IMAGE clone 646675, which encodes EST AA205581, and a human fetal brain cDNA from Clonetech. The micleotide sequence of a cDNA encoding a partial ADAMTS-9 mouse

protein was obtained using IMAGE clone 535663, which encodes EST AA1 06215, and a mouse cDNA library obtained from Clonetech. RACE was performed as described above in Example 1. The nucleotide sequence of the cDNA encoding the full-length ADAMTS-9 human proteinand the amino acid sequence of such protein is shown in Fig.6. The nucleotide sequence of the cDNA encoding the partial ADAMTS-9 mouse protein and the amino acid sequence of such protein is shown in Fig. 7.

The predicted Mr of the mature human ADAMTS-9 protein is 189777.20 daltons. The prodomain of the predicted ADAMTS-9 protein has 3 consensus cleavage signal for furin. The catalytic domain of the ADAMTS-9 contains eight cysteine residues and the reprolysin - zinc binding signature sequence, HELGHVFNMPHD, which is followed by a "Met-turn". The catalytic domain is followed by a domain with 25-30% similarity to snake venom disintegrins The disintegrin domain contains eight cysteine residues. The first TS repeat contains is followed by a conserved CRD sequence which. contains 10 conserved cysteines. The spacer domain of ADAMTS-9 is 124 amino acids in length and is followed by 14 additional TS modules and a C-terminal domain. The ADAMTS-9 protein contains 6 potential glycosylation sites within the mature protease:, one in the spacer domain, one in TSP 1 -7, one in TSPI-8, and 3 in the C-terminal domain. The ADAMTS-9 bears 44% sequence identity to ADAMTS-4.

Example 6: ADAMTS-10

The nucleotide sequence of a cDNA encoding a fall-length ADAMTS- 10 protein was obtained using IMAGE clone 110403, which encodes EST AA588434, and a human fetal brain cDNA from Clonetech. The nucleotide sequence of a cDNA encoding a partial, mouse ADAMTS-10 protein was obtained using IMAGE clone 1077653, which encodes EST AA822090, and a mouse embryo cDNA library from Clon tech. RACE was

performed as described above in Example 1. The nucleotide sequence of the human ADAMTS-10 cDNA and the predicted amino acid sequence, SEQ ID 18, of the human ADAMTS-10 protein encoded by such DNA is shown in Fig. 9. The nucleotide sequence of the cDNA encoding the partial mouse ADAMTS-10 protein and the amino acid sequence of such protein is shown in Fig. 10.

The predicted Mr of the mature ADAMTS-10 protein is 95238 daltons. The pro-domain of the full-length ADAMTS-10 protein has no consensus cleavage signal for furin. The catalytic domain of the ADAMTS-10 contains eight cysteine residues and the reprolysin-zinc binding signature sequence, HEIGHTFGMNHD, which is followed by a "Met-turn". The catalytic domain is followed by a domain with 30% similarity to snake venom disintegrins. The disintegrin-like domain contains eight cysteine residues. The first TS repeat is followed by a conserved CRD sequence which contains 8 conserved cysteines. The spacer domain of ADAMTS-10 is followed by 4 additional TS modules and a Kunitz domain. The ADAMTS-10 protein contains 2 potential glycosylation sites within the mature protease: one in the catalytic domain, and one in the TS 1-3 domain. ADAMTS-10 bears approximately 40% sequence identity to ADAM-TS1, which is characterized as being involved in inflammation.

Comparison of the ADAMTS-N Proteins.

As shown in Figure 11, the ADAMTS-5. ADAMTS-6, and ADAMTS-7 proteins share a common domain organization. From amino to carboxyl termini, they are as follows:

1. **A pre-pro region.** A typical signal sequence of variable length is followed by a putative pro-region of variable length but demonstrating short stretches of sequence identity. Three cysteine residues are, predicted within each novel pro-domain, of which the most C-terminal is an "asymmetric" cysteine lying within a sequence

context similar to the cysteine "switch" of the MMPs. All three novel cDNAs predict consensus cleavage signals for furin, three in the case of ADAMTS-5, and one each in the case of ADAMTS-6 and ADAMTS-7. The most carboxyl-terminal furin cleavage site in ADAMTS-5 predicts the processing site for generation of the mature protease. The amino terminus of the mature proteins is predicted to start at the residue immediately following the cleavage sites.

2. **A catalytic domain.** The catalytic domains are very similar to each other and contain eight cysteine residues and a typical reprolysin-type zinc binding signature followed by a "Met-turn". Five cysteine residues are upstream of the zinc binding sequence, while three residues are downstream, an arrangement that is shared with other ADAMTS members. The methionine of the met-turn is not at a constant distance from the zinc-binding signature, but in all three novel proteases, a constant cysteine residue is present in that interval.

3. **A disintegrin-like domain.** The catalytic domain is followed by a domain of 60-90 residues with 35-45% similarity to snake venom disintegrins, but without the canonical cysteine arrangement seen in the latter. This disintegrin-like domain is of comparable length in ADAMTS-5 and ADAMTS-7, it is considerably shorter in ADAMTS-6.

4. **A TS module.** The first TS repeat is very similar in all three novel proteases and very similar to the first TS repeat of other ADAMTSs. It contains the same number of residues (fifty-two) in all three novel proteins.

5. **The cysteine-rich domain.** This TS domain is followed by a conserved cysteine-rich sequence termed the cysteine-rich domain (CRD).

6. **The spacer domain.** This domain is of variable length, in all ADAMTSs and lacks the sequence landmarks so characteristic of all the

other domains. It shows the least homology of all the domains.

7. **A C-terminal TS module.** The sequence of the second TS module is more variant between the members of the ADAMTS family than the first TS module, despite the conservation of the number and spacing of cysteine residues.

Overall, the predicted mature forms of these proteases show 20-30% similarity to each other and to ADAMTS1-4 although this may be considerably higher or lower for individual domains as described above.

ADAM-TS9 and ADAM-TS10 contain all the domains present in ADAMTS-5 through ADAMTS-8. In addition, ADAMTS-9 and ADAMTS-10 contain the following domains:

A. ADAMTS-9: After the c-terminal TS1 domain which is present in ADAMTS5-8, ADAMTS-9 contains 13 additional and homologous TS11 domains, thus, ADAMTS-9 contains a total of 15 TS1 domains, of which 14 are adjacent to each other in the c-terminal half of the molecule. The 15th TS1 domain from the N-terminus is followed by a unique c-terminal domain which does not possess recognizable domain structure and contains 196 residues including 9 cysteine residues.

B. ADAMTS-10: After the c-terminal TS1 domain which is present in ADAMTS 8, ADAMTS-10 contains 3 additional and homologous TS1 domains, thus, that ADAMTS-10 contains a total of 5 TS1 domains, of which 4 are adjacent to each other in the c-terminal half of the molecule. The 5th TS 1 domain from the N-terminus is followed by an additional 47 amino acid residues including six (6) cysteine residues. These 47 residues have sequence similarity of 30%-40% to the c-terminus of pro-hormone convertase 5 and 6, and to the Kunitz family of inhibitors.

Northern Analysis

Mouse embryo northern blots and multiple tissue northern blots

from human and mouse tissues (Clontech, Palo Alto, CA) were hybridized to the [$\alpha^{32}P$]-dCTP labeled inserts of I.M.A.G.E. clones as per the manufacturer's recommendations followed by autoradiographic exposure for 3-7 days.

In situ hybridization used cryosections of mouse embryos of gestational age 8.5 days and 10.5 days. Embryos were collected with the inclusion of the surrounding uterus and fixed overnight in 4% paraformaldehyde. Sense and anti-sense probes continuously labeled with digoxigenin-UTP (Boehringer-Mannheim, Indianapolis, IN) were transcribed with T7 and T3 RNA polymerases, respectively, using as template a 63 0 bp EcoRI-Sacl fragment from the *Adamts*-5 clone 569515 (Fig. 14) cloned into pBluescript SK+ (Stratagene, La Jolla, CA). In situ hybridization was done essentially as previously described in Apte, et al. (1997) J. Biol. Chem. 272:2551-25517, which is specifically incorporated herein by reference, except that sections were predigested with proteinase K (Boehringer-Mannheim, Indianapolis, IN) at a lower, concentration (1 -5 μg/ml) than reported in Apte, et al.. Bound, digoxigenin-labeled probe was detected using an alkaline phosphatase tagged anti-digoxigenin antibody (Boehringer-Mannheim, Indianapolis, IN) and nuclei were counterstained with methyl green.

Specific hybridization of the antisense *Adamts*-5 probe to sections of 8.5 day-old mouse embryos was obtained, whereas only low background staining was noted with the control sense probe. Staining was uniform throughout the 8.5 day old embryos. In addition, there was labeling of mRNA in trophoblastic cells lining the uterine cavity as well as in the developing placenta (Fig. 14). The decidual reaction within the uterus also showed upregulation of *Adamts*-5 mRNA relativ to the negative controls, In sections from 10.5 day old embryos, labeling was widespread but less intense compared to the 8.5

day-old embryo. Labeled cells were seen in mesenchyme and somites as well as in the neural tube and developing hindgut. Northern analysis also indicated that mRNA encoding ADAMTS-5 was present in human placenta but was barely detectable in adult lung, heart, brain, liver, skeletal muscle, kidney and pancreas.

Northern analysis showed undetectable expression of *Adamts*-6 during mouse embryo development. Northern analysis indicated that mRNA encoding ADAMTS-6 was present in human placenta but was barely detectable in adult lung, heart, brain, liver, skeletal muscle, kidney and pancreas. *Adamts*-7 was expressed at low levels throughout mouse development. In adult human tissues examined with human cDNA probes, ADAMTS-7 mRNA was found in all tissues examined, i.e. in lung, heart, brain, liver, skeletal muscle, kidney, pancreas and placenta. The sizes of the mRNA species recognized by the probes varied. ADAMTS-5 mRNA was approximately 10 kbp in size in human tissue. The most prominent *Adamts*-5 species was estimated at 7.5 kbp together with additional bands at 10 kbp and 4.5 kbp. The lone mRNA species detected by ADAMTS-6 probe was approximately 8.5 kbp, whereas the most common mRNA species detected by ADAMTS-7 probe 5 was 5 kbp in size with an additional species seen at 7 kbp in skeletal muscle.

In mouse, ADAMTS-8 is expressed during fetal development (days 7, 11, 15, 17) and in adult mouse lung and heart with an mRNA size of approximately 3.8 kbp. In adult human tissue, ADAMTS-8 is expressed in lung and brain but not in heart, muscle, kidney, colon or thymus. The mRNA size is 3.8 kbp.

ADAMTS-9 is expressed in lung, ovary placenta, heart, brain, muscle, kidney and pancreas with a mRNA size of 8 kb. In addition, kidney and ovary contain additional transcripts of size 3 kb and 4.4 kb respectively. These additional transcripts may represent alternatively spliced or short forms of ADAMTS9.

ADAMTS-10 is expressed in thymus, prostate, testis, ovary, small intestine, colon, peripheral blood leukocytes, heart, brain, placenta, lung, liver, muscle, kidney and pancreas, as well as in many cell lines such as A549, HeLa and K562. There are two transcripts of 5 kb and 8kb present in all tissues.

Example 7: ADAMTS-R1

The nucleotide sequence of a cDNA encoding a full-length ADAMTS-R1 protein was obtained using IMAGE clone 752797 which encodes EST AA, and a human fetal brain cDNA from Clontech. RACE was performed as described above in Example 1. The nucleotide sequence, SEQ ID NO:21, of the ADAMTS-R1 cDNA and the predicted amino acid sequence, SEQ ID NO:22, of the ADAMTS-R1 protein encoded by such DNA is shown in Fig. 11.

The predicted Mr of the full-length, unprocessed ADAMTS-R1 protein is 58358.20 daltons. The domain organization of the ADAMTS-10 protein is shown in Fig. 15. In contrast to the ADAMTS-N proteins of examples 1-6, ADAMTS-R1 protein does not have a pro-metalloprotease or disintegrin-like domain or a consensus cleavage signal for furin. ADAMTS-R1 has a signal(pre) peptide which is followed by a first TS module and a conserved CRD sequence which contains 10 conserved cysteines. The spacer domain of ADAMTS-R1 is 115 amino acids in length and is followed by 3 additional TS modules and a short sequence of 33 amino acids. The ADAMTS-R1 protein contains one potential glycosylation sites which is in the spacer domain. ADAMTS-R1 bears 30-40% sequence identity to ADAMTS1 and ADAMTS4 in the related domains. ADAMTS-R1 mRNA is present in human heart, brain, kidney, muscle, lung, placenta, testis, ovary, colon, intestine, and prostate. There are three transcripts of 2.5 kb, 4.7 kb and 6.5 kbp present in all such tissues. In mouse, expression is seen in skeletal muscle, and the transcript size is 6.5 kb.

Although certain embodiments of this invention have been shown and described, various adaptations and modifications can be made without departing from the scope of the invention as defined in the appended claims.

CLAIMS

1. An isolated mammalian protein selected from the group consisting of an ADAMTS-5 protein an ADAMTS-6 protein, an ADAMTS-7 protein, an ADAMTS-8 protein, an ADAMTS-9 protein, an ADAMTS-10 protein, and an ADAMTS-R1 protein.

2. The isolated mammalian protein of claim 1 wherein said protein comprises an amino acid sequence which is at least 95% identical to a sequence selected from the group consisting of: amino acid 262 through amino acid 930 of SEQ ID NO:2; amino acid 1 through amino acid 518 of SEQ ID NO:4; amino acid 245 through amino acid 860 of SEQ ID NO:6; amino acid 233 through amino acid 997 of SEQ ID NO:8; amino acid 229 through amino acid 905 of SEQ ID NO:10; amino acid 1 through amino acid 245 of SEQ ID NO:12; amino acid 236 through amino acid 1882 of SEQ ID NO:14; amino acid 1 through amino acid 874 of SEQ ID NO:16; amino acid 212 through amino acid 1081 of SEQ ID NO:18; amino acid 1 through amino acid 450 of SEQ ID NO:20; and amino acid 1 through amino acid 547 of SEQ ID NO:22.

3. The isolated protein of claim 2 wherein said amino acid sequence further comprises a prepropeptide sequence at the amino terminus thereof.

4. The isolated protein of claim 1 wherein said protein is a human ADAMTS-5 protein or a mouse ADAMTS-5 protein.

5. The isolated protein of claim 1 wherein said protein is a human ADAMTS-6 protein.

6. The isolated protein of claim 1 wherein said protein is a human ADAMTS-7 protein.

7. The isolated protein of claim 1 wherein said protein is a mouse ADAMTS-8 or a human ADAMTS-8 protein.

8. The isolated protein of claim 1 wherein said protein is a human

ADAMTS-9 or a mouse ADAMTS-9 protein.

9. The isolated protein of claim 1 wherein said protein is a human ADAMTS-10 or a mouse ADAMTS-10 protein.

10. The isolated protein of claim 1 wherein said protein is a human ADAMTS-R1 protein.

11. An isolated polynucleotide comprising a sequence which encodes a mammalian protein selected from the group consisting of an ADAMTS-5 protein, an ADAMTS-6 protein, an ADAMTS-7 protein, an ADAMTS-8 protein, an ADAMTS-9 protein, an ADAMTS-10 protein, and an ADAMTS-R1 protein.

12. The isolated polynucleotide of claim 11 wherein said protein comprises an amino acid sequence which is at least 95% identical to a sequence selected from the group consisting of: amino acid 262 through amino acid 930 of SEQ ID NO:2; amino acid 1 through amino acid 518 of SEQ ID NO:4; amino acid 245 through amino acid 860 of SEQ ID NO:6; amino acid 233 through amino acid 997 of SEQ ID NO:8; amino acid 229 through amino acid 905 of SEQ ID NO:10; amino acid 1 through amino acid 245 of SEQ ID NO:12; amino acid 236 through amino acid 1882 of SEQ ID NO:14; amino acid 1 through amino acid 874 of SEQ ID NO:16; amino acid 212 through amino acid 1081 of SEQ ID NO:18; amino acid 1 through amino acid 450 of SEQ ID NO:20, and amino acid 1 through amino acid 547 of SEQ ID NO:22.

13. The isolated polynucleotide of claim 11 wherein said nucleotide sequence encodes a protein having a signal sequence at the amino terminus thereof.

14. The isolated polynucleotide of claim 11 wherein said polynucleotide comprises a sequence selected from the group consisting of: nucleotide 800 through nucleotide 2810 of SEQ ID NO:1 of an all lic variant thereof; nucleotide 1 through

nucleotide 1519 of SEQ ID NO:3 or an allelic variant thereof; nucleotide 754 through nucleotide 2602 of SEQ ID NO:5 or an allelic variant thereof; nucleotide 708 through nucleotide 3003 of SEQ ID NO:7 or an allelic variant thereof; nucleotide 962 through nucleotide 2992 of SEQ ID NO:9 or an allelic variant thereof; nucleotide 1 through nucleotide 739 of SEQ ID NO:11 or an allelic variant thereof; nucleotide 708 through nucleotide 5648 of SEQ ID NO:13 or an allelic variant thereof; nucleotide 1 through nucleotide 2625 of SEQ ID NO:15 or an allelic variant thereof; nucleotide 634 through nucleotide 3243 of SEQ ID NO:17 or an allelic variant thereof; nucleotide 1 through nucleotide 1642 of SEQ ID NO:19 or an allelic variant thereof; and nucleotide 51 through nucleotide 1625 of SEQ ID NO:21 or an allelic variant thereof.

15. The isolated polynucleotide of claim 11 wherein said polynucleotide hybridizes under stringent conditions to a nucleic acid molecule comprising a sequence complementary to the protein encoding sequence of SEQ ID NO:1; SEQ ID NO:3; SEQ ID NO:5; SEQ ID NO:7; SEQ ID NO:9; SEQ ID NO:11; SEQ ID NO:13; SEQ ID NO:15; SEQ ID NO:17; SEQ ID NO:19; or SEQ ID NO:21.

16. An isolated polynucleotide having a sequence which is complementary to the protein encoding sequence of the polynucleotide of claim 11.

17. An expression vector comprising a polynucleotide of claim 11.

18. A host cell transformed or transfected with an expression vector of claim 17.

19. A method for producing an ADAMTS-N protein or an ADAMTS-R1 protein, said method comprising the steps of

(a) culturing a host c ll of claim 18 under conditions suitable for expr ssion of an ADAMTS-N protein or an ADAMTS-R1

protein; and

(b) recovering said ADAMTS-N protein or said ADAMTS-R1 protein from the host cell culture.

20. An antibody that binds to a protein selected from the group consisting of an ADAMTS-5 protein, an ADAMTS-6 protein, an ADAMTS-7 protein, an ADAMTS-8 protein, an ADAMTS-9 protein, an ADAMTS-10 protein and an ADAMTS-R1 protein.

21. An oligopeptide for producing an antibody that binds to an ADAMTS-N protein or an ADAMTS-R1 protein wherein said oligopeptide has a sequence selected from the group consisting of:

a) SVSIERFVETLVVADK, SEQ ID NO:23;

b) EVAEAANFLALRSEDPDKY, SEQ ID NO:24;

c) VKEDVENPKAVVDGDWGP, SEQ ID NO:25;

d) QHPFQNEDYRPRSASPSRTH, SEQ ID NO:26;

e) PQNCKEVKRLKGASEDGEYF, SEQ ID NO:27;

f) QELEEGAAVSEEPS, SEQ ID NO:28;

g) YYPENIKPKPKLQE; SEQ ID NO:29;

h) HIKVRQFKAKDQTRF; and

i) CEAKNGYQSDAKGVKTFVEWVPKYAG, SEQ ID NO:30.